

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(43) 国际公布日:
2005年6月23日(23.06.2005)

PCT

(10) 国际公布号:
WO 2005/056591 A1

(51) 国际分类号⁷: C07K 14/47, C12N 15/12, 15/63, 15/10, A61K 38/17, A61P 31/04

(21) 国际申请号: PCT/CN2004/001435

(22) 国际申请日: 2004年12月10日(10.12.2004)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
200310117354.6 2003年12月11日(11.12.2003) CN

(71) 申请人(对除美国以外的所有指定国): 中国人民解放军军事医学科学院放射医学研究所(BEIJING INSTITUTE OF RADIATION MEDICINE) [CN/CN]; 中国北京市海淀区太平路27号, Beijing 100850 (CN)。

(72) 发明人: 及

(75) 发明人/申请人(仅对美国): 赵士富(ZHAO, Shifu) [CN/CN]; 张咏(ZHANG, Yong) [CN/CN]; 尹丽莉(YIN, Lili) [CN/CN]; 中国北京市海淀区太平路27号, Beijing 100850 (CN)。

(74) 代理人: 北京纪凯知识产权代理有限公司(JEEKAI & PARTNERS); 中国北京市西城区宣武门西大街甲129号金隅大厦602室, Beijing 100031 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

本国际公布:
— 包括国际检索报告。

所引用双字母代码和其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: GLYCINE-RICH PROTEINS, THEIR CODING SEQUENCES AND APPLICATIONS

(54) 发明名称: 富含甘氨酸蛋白及其编码基因与应用

(57) Abstract: The present invention relates to the glycine-rich proteins, their coding sequences and applications, especially to the glycine-rich proteins, their coding sequences and antibacterial applications. The glycine-rich proteins of the present invention are at least one selected from the following protein family: the proteins having the sequence of SEQ ID NO: 1, SEQ ID NO: 3 to SEQ ID NO: 14, or the proteins having the sequence by making deletion, insertion, replacement and addition of amino acid groups to the sequence of SEQ ID NO: 1 and SEQ ID NO: 3 to SEQ ID NO: 14 and having antibacterial activity. The glycine-rich proteins are useful for antibacterium, such as useful for preparation of preventive and therapeutically medicament for bacterial infection diseases of human or livestock, for producing the different biology preparation for preventing and treating potential bacterial infection, for producing transgenosis antimicrobial organism, for producing the derivatives, agonists, ligands and antibodies of the mentioned glycine-rich proteins.

(57) 摘要

本发明涉及富含甘氨酸蛋白及其编码基因与应用, 特别涉及富含甘氨酸蛋白及其编码基因与它们在抗细菌中的应用。本发明的富含甘氨酸蛋白, 选自下述蛋白质家族中的至少一种: 具有序列表中序列1、序列3-14的氨基酸残基序列的蛋白质或将序列表中序列1、序列3-14的氨基酸残基序列经过1至20个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加1至20个氨基酸残基且具有抗细菌作用的蛋白质。本发明的富含甘氨酸蛋白及其编码基因可用于抗细菌, 如用于制备预防和/或治疗人或畜细菌性感染疾病的药物; 用于制备预防和/或治疗具有潜在细菌感染的不同种类生物用制品; 用于生产抗病虫害的转基因生物; 用于制备针对所述富含甘氨酸蛋白的衍生物或拮抗剂及其配体、抗体。

富含甘氨酸蛋白及其编码基因与应用

技术领域

- 本发明涉及富含甘氨酸蛋白及其编码基因与应用，特别涉及人和小鼠
- 5 富含甘氨酸蛋白及其编码基因，及它们在在抗细菌中的应用。

背景技术

- 20 多年来，医药界没有发现任何新的抗菌素家族，但是抗药性现象的发展却极为迅速，人们甚至可以用肉眼观察到细菌的基因突变。更为可怕的是，对付金黄色葡萄球菌的障碍刚刚攻克，美国医生就发现了对可有效
- 10 抑制细菌的最新抗菌素—万古霉素具有抗性的菌株。致病菌的抗药性问题已经日益严重地威胁着人们的健康。寻找全新类型的抗生素是解决抗药性问题的一条有效途径。抗菌肽作为生物天然免疫的活性分子，被各种生物用来抵御来自外界病菌的侵染，在生物界中广泛存在。人为分离出来的天然抗菌肽表现出抗菌活性高，抗菌谱广，种类多，可供选择的范围广，靶
- 15 菌株不易产生抗性突变等特点，因而在医药工业、食品工业和农业上被认为有着广阔的应用前景。

- 抗菌肽是指由基因编码，在核糖体上合成的相对分子质量通常在 10kDa 以下，具有抗菌活性的多肽类物质，也叫多肽抗生素。大部分抗菌肽具有热稳定性，在 100℃ 下加热 10~15min 仍能保持其活性。抗菌肽对较大的
- 20 离子强度和较高或较低的 pH 值均具有较强的抗性。多数抗菌肽的等电点大于 7，表现出较强的阳离子特征。同时，部分抗菌肽尚具备抵抗胰蛋白酶或胃蛋白酶水解的能力。此外，研究发现不同家族的抗菌肽之间在序列上极少存在同源性，而同一家族的不同成员之间序列上却存在高度的保守性，这意味着其功能也非常保守。抗菌肽除了具有抗细菌或真菌的作用外，
- 25 有些还具有抗原虫、病毒或癌细胞的功能。抗菌肽在生物界中广泛存在。迄今为止，人们已经从细菌、真菌，到两栖类、昆虫、高等植物、哺乳动物、直至人类体内发现了多达 700 种以上的多肽抗生素。目前，已有多种抗菌肽正在进行临床前的可行性研究；此外，抗菌肽转基因动物、转基因植物以及抗菌肽在食品防腐，鲜花保鲜，化妆品，种子包衣和动物饲料添加
- 30 剂等方面的应用研究也正在进之中。

鉴于抗菌肽在医药工业、食品工业和农业上有着广阔的应用前景，通过各种途径从生物界中寻找新型的天然抗菌肽已成为目前世界范围内热点之一。

发明公开

- 5 本发明的目的是提供一种富含甘氨酸蛋白及其编码基因以及它们的抗菌用途。

本发明所提供的富含甘氨酸蛋白，名称为Glyrichin，选自（克隆自）下述蛋白质家族（Glyrichin家族）中的至少一种：

- 1) 人 Glyrichin (hGlyrichin) 和小鼠 Glyrichin (mGlyrichin) :
10 均具有序列表中序列 1 的氨基酸残基序列的蛋白质或将序列表中序列 1 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

- 2) 斑马鱼 (*Danio rerio*) Glyrichin: 具有序列表中序列 3 的氨基酸
15 残基序列的蛋白质或将序列表中序列 3 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

- 3) 按蚊 (*Anopheles gambiae*) Glyrichin: 具有序列表中序列 4 的氨基酸残基序列的蛋白质或将序列表中序列 4 的氨基酸残基序列经过 1 至 20
20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

- 4) 果蝇 (*Drosophila melanogaster*) Glyrichin: 具有序列表中序列 5
25 的氨基酸残基序列的蛋白质或将序列表中序列 5 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

- 5) 线虫 (*Caenorhabditis elegans*) Glyrichin: 具有序列表中序列
6 的氨基酸残基序列的蛋白质或将序列表中序列 6 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

- 30 6) 线虫 (*Caenorhabditis elegans*) Glyrichin: 具有序列表中序列

7 的氨基酸残基序列的蛋白质或将序列表中序列 7 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

5 7) 芽殖酵母 (*Schizosaccharomyces pombe*) Glyrichin: 具有序列表中序列 8 的氨基酸残基序列的蛋白质或将序列表中序列 8 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

8) 酿酒酵母 (*Saccharyomyces cerevisiae*) Glyrichin: 具有序列表中序列 9 的氨基酸残基序列的蛋白质或将序列表中序列 9 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

9) 拟南芥 (*Arabidopsis thaliana*) Glyrichin: 具有序列表中序列 10 的氨基酸残基序列的蛋白质或将序列表中序列 10 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

10) 疟原虫 (*Plasmodium falciparum* 3D7) Glyrichin: 具有序列表中序列 11 的氨基酸残基序列的蛋白质或将序列表中序列 11 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

20 11) 疟原虫 (*Plasmodium yoelii yoelii*) Glyrichin: 具有序列表中序列 12 的氨基酸残基序列的蛋白质或将序列表中序列 12 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

12) 稻瘟病菌 (*Magnaporthe grisea*) Glyrichin: 具有序列表中序列 13 的氨基酸残基序列的蛋白质或将序列表中序列 13 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

13) 脉孢菌 (*Neurospora crassa*) Glyrichin: 具有序列表中序列 14 的氨基酸残基序列的蛋白质或将序列表中序列 14 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末

端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质。

所述富含甘氨酸蛋白优选为人 Glyrichin 和小鼠 Glyrichin。

其中，所述缺失、插入和/或取代以及在羧基末端和/或氨基末端添加的氨基酸残基的数目优选为 1 至 10 个，更有选为 1 至 5 个，最优选为 1 至 3 个。

所述取代方式优选为如表 1 所示的取代。

表 1. Glyrichin 蛋白家族可取代氨基酸列表

| 最初的残基 | 代表性的取代 | 优选的取代 |
|---------|-------------------------|-------|
| Ala (A) | Val; Leu; Ile | Val |
| Arg (R) | Lys; Gln; Asn | Lys |
| Asn (N) | Gln; His; Lys; Arg | Gln |
| Asp (D) | Glu | Glu |
| Cys (C) | Ser | Ser |
| Gln (Q) | Asn | Asn |
| Glu (E) | Asp | Asp |
| Gly (G) | Pro; Ala | Ala |
| His (H) | Asn; Gln; Lys; Arg | Arg |
| Ile (I) | Leu; Val; Met; Ala; Phe | Leu |
| Leu (L) | Ile; Val; Met; Ala; Phe | Ile |
| Lys (K) | Arg; Gln; Asn | Arg |
| Met (M) | Leu; Phe; Ile | Leu |
| Phe (F) | Leu; Val; Ile; Ala; Tyr | Leu |
| Pro (P) | Ala | Ala |
| Ser (S) | Thr | Thr |
| Thr (T) | Ser | Ser |
| Trp (W) | Tyr; Phe | Tyr |
| Tyr (Y) | Trp; Phe; Thr; Ser | Phe |
| Val (V) | Ile; Leu; Met; Phe; Ala | Leu |

本发明的富含甘氨酸蛋白可进行修饰（通常不改变一级结构），如乙酰化，羧基化，糖基化和磷酸化。

上述富含甘氨酸蛋白的编码基因也属于本发明的保护范围。

其中，所述人 Glyrichin (hGlyrichin) 的编码基因可具有序列表中
5 SEQ ID No: 2 的 DNA 序列或与序列表中 SEQ ID No: 2 限定的 DNA 序列具有 90% 以上同源性，且编码序列表中序列 1 的氨基酸残基序列的 DNA 序列或在高严谨条件下可与序列表中的序列 2 限定的 DNA 序列杂交的核苷酸序列；

所述高严谨条件为杂交系指用含 $0.1 \times$ SSPE (或 $0.1 \times$ SSC)、0.1% SDS
10 的溶液在 65°C 下洗膜。

含有上述富含甘氨酸蛋白的编码基因的表达载体，细胞系及工程菌均属于本发明的保护范围。

本发明所述的富含甘氨酸蛋白可用于抗细菌，具体包括以下几方面的用途：

- 15 1) 用于制备预防和/或治疗人或畜细菌性感染疾病的药物。
 2) 用于制备预防和/或治疗具有潜在细菌感染的不同种类生物用制品。
 3) 用于生产抗病虫害的转基因生物。
 4) 用于制备针对所述富含甘氨酸蛋白的衍生物或拮抗剂及其配体、
20 抗体。

附图说明

图 1 为 Glyrichin 家族成员同源性比较中的 12 条序列进化树分析图

图 2 为保守的 Glyrich 结构域

图 3 为 Glyrich 结构域在不同成员间的序列一致性

25 图 4 为 hGlyrichin 家族的结构特点图

图 5A 为以 hGlyrichin 基因为探针检测 hGlyrichin 转录本大小和表达谱的 Northern blot 分析

图 5B 为人 Glyrichin 体外转录和翻译试验的结果

30 图 6A 为转化了 pET-22b(+) 空载体的大肠杆菌 BL21 在 0.5mM IPTG 诱导和不诱导条件下的半对数生长曲线

图 6B 为转化了 pET-22b (+) -UBF 和 pET-22b (+) -PTP 的大肠杆菌 BL21 在 0.5mM IPTG 诱导和不诱导条件下的半对数生长曲线

图 6C 为转化了 pET-22b-hGlyrichin 阳性克隆 1 和 8 的大肠杆菌 BL21 在 0.5mM IPTG 诱导和不诱导条件下的半对数生长曲线

5 图 6D 为转化了 pET-22b-hGlyrichin 阳性克隆 1 的大肠杆菌 BL21 在 0.5mM IPTG 诱导和不诱导条件下的两次重复实验结果的半对数生长曲线

图 7 为 hGlyrichin 原核表达产物的 PAGE 电泳结果

图 8 为全长 hGlyrichin 和 5' -缺失体基因转化酵母后表达产物对大肠杆菌 BL-21 菌体生长的抑制作用

10 图 9 为 hGlyrichin 全长、5' 和 3' 缺失体基因对转化后 BL-21 菌生长的影响

实施发明的最佳方式

发明人利用抑制性差减杂交的方法，对 LTC (long term culture) 培养前、后的小鼠骨髓基质细胞差异表达基因进行了大量筛选，获得了 131
15 个差异表达的 EST 克隆。生物信息学分析证实，这些克隆代表了 26 种已知或部分已知功能的基因和 7 条全新的基因，其中 5 条具有完整的开放阅读框架，鼠源的 mGlyrichin 基因 (GenBank 号为 AY028425) 就是其中之一，编码序列具有表中序列 1 的氨基酸残基序列的富含甘氨酸蛋白 mGlyrichin。

20 在得到鼠源 mGlyrichin 后，发明人根据鼠源 Glyrichin 基因设计了一对引物，以人胎肝 mRNA 为模板，通过 RT-PCR 扩增获得人源 Glyrichin 完整的 ORF 序列共 240bp，以及根据 ORF 序列推测的氨基酸序列。发明人利用生物信息学手段对序列进行了详细分析，进行了 Glyrichin 家族成员间的同源性比较—多序列对齐。

25 为研究 Glyrichin 家族各成员的亲缘关系，对图 1 所示的 12 条序列进行了进化树分析，图 1 显示出 Glyrichin 基因从低等到高等的进化历程及各成员间亲缘关系的远近。Glyrichin 家族成员间保守的甘氨酸富含区 (Glyrich domain) 如图 2 所示 (图中，保守的氨基酸用反显字表示)。富含甘氨酸结构域 (Glyrich 结构域) 在 Glyrichin 家族不同成员间的序列一致性如图 3 所示，可以看出每两个成员间的亲缘关系的远近。电子 PCR
30

结果显示, hGlyrichin 基因定位于人染色体 20q11.21 区域, 由三个外显子组成, 编码 79 个氨基酸的小分子蛋白。氨基酸水平 BlastP 同源性比较显示该基因在进化上十分保守, 存在于从真菌(如脉孢菌、酵母菌等)、植物(如拟南芥等)、疟原虫、线虫、果蝇、按蚊、斑马鱼、小鼠到人等物种中, 组成 Glyrichin 同源性蛋白家族; 在不同种属间显示出很高的同源性, 已知人源 Glyrichin 与鼠源相比 100%同源, 与斑马鱼 Glyrichin 相比具有 90%同源性, 与果蝇的同源性有 62%, 与酵母的同源性也有 46%; 家族内不同成员间的亲缘关系以进化树表示(图 1)。

其中, 具有序列 1、序列 3-14 的氨基酸残基序列的 Glyrichin 蛋白家族成员的序列特征如表 2 所示。

表 2. 具有序列 1、序列 3-14 的氨基酸残基序列的 Glyrichin 蛋白家族成员的序列特征

| 物种属名 | 多肽长度 (aa) | 等 电 点 | 电荷 数(pH7.0) | 甘氨酸含 量(%) |
|----------------------------------|--------------|----------|----------------|--------------|
| <i>Homo sapiens</i> | 79 | 9.36 | 4.79 | 21.52 |
| <i>Mus musculus</i> | 79 | 9.36 | 4.79 | 21.52 |
| <i>Danio rerio</i> | 80 | 9.81 | 4.82 | 21.2 |
| <i>Anopheles gambiae</i> | 128 | 9.73 | 10.08 | 14.1 |
| <i>Drosophila melanogaster</i> | 79 | 9.66 | 4.82 | 20.3 |
| <i>Caenorhabditis elegans</i> | 145 | 10.79 | 12.01 | 14.48 |
| <i>Schizosaccharomyces pombe</i> | 107 | 10.04 | 4.05 | 13.08 |
| <i>Saccharomyces cerevisiae</i> | 112 | 9.74 | 5.04 | 13.27 |
| <i>Plasmodium falciparum</i> 3D7 | 168 | 9.66 | 5.15 | 15.94 |
| <i>Plasmodium Yoelii Yoelii</i> | 167 | 9.16 | 4.15 | 15.28 |
| <i>Magnaporthe grisea</i> | 107 | 11.75 | 8.24 | 13.08 |
| <i>Arabidopsis thaliana</i> | 74 | 9.84 | 6.18 | 14.86 |

其中，人 Glyrichin (hGlyrichin) 由 79 个氨基酸残基组成，分子量为 8.8kDa，等电点 9.36，在 pH 值为 7 的条件下带 4.79 个正电荷，甘氨酸含量为 21.52%。发明人进行了 hGlyrichin 二级结构预测，hGlyrichin 的结构特点如图 4，为 SignalP 分析结果。Glyrichin 家族其他成员的等电点均大于 7，在 pH=7 时都带正电荷，且都存在强疏水区域。上述结果表明 Glyrichin 家族具有已知抗菌肽典型的结构特征。

本发明的 13 个 Glyrichin 蛋白家族成员分布在从真菌到人的所有已完成基因组测序的物种中。本领域的普通技术人员可以轻易地据此推断出它也可能存在于其他基因组尚未完成测序的物种中。结构上共同的特点是存在一个甘氨酸 (Glycine) 富含区，长度在 59-68 个氨基酸之间。该区域基本涵盖了人源和鼠源 Glyrichin 全部的序列。存在于其他物种中的 Glyrichin 除了共同的甘氨酸富含区外，还在此保守区的氨基端或羧基端存在其他序列 (如信号肽等)，但不具有保守性 (图 2、图 3)。这说明甘氨酸保守区决定了 Glyrichin 蛋白家族成员最基本的生物活性；也从另一方面说明生物在进化过程中逐渐淘汰了冗余序列，使分子变得更短小、结构更紧凑。

通过常规的重组 DNA 技术，可利用本发明的富含甘氨酸蛋白的编码基因表达或生产重组的富含甘氨酸蛋白。如可按以下方法进行表达或生产：

- (1) 用含有上述富含甘氨酸蛋白编码基因的重组表达载体转化或转导宿主细胞；
- (2) 培养所述宿主细胞；
- (3) 从培养基或细胞中分离、纯化蛋白质。

用于构建所述重组表达载体的出发载体可为本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒如腺病毒或其它载体。总之，只要能在宿主体内复制和稳定表达质粒和载体都可以用。

表达载体的一个重要特征是通常含有复制起点、启动子、标记基因和翻译控制元件。本领域的技术人员熟知的方法能用于构建含 Glyrichin 编码 DNA 序列和合适的转录/翻译控制信号的表达载体。

本领域普通技术人员都清楚如何选择适当的载体、启动子、增强子和

宿主细胞。用所述重组表达载体转化宿主细胞可用本领域技术人员熟知的常规技术进行。获得的转化子可用常规方法培养，表达本发明的基因编码的富含甘氨酸蛋白。在上述方法中的富含甘氨酸蛋白可在细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要，可利用其物理的、化学的和其
5 它特性通过各种分离方法分离和纯化重组的蛋白。在使用本发明 Glyrichin 蛋白时，还可同时使用其它药剂，如青霉素等抗菌素。

在本发明中，术语“Glyrichin ”、“Glyrichin 家族成员”、“Glyrichin 蛋白”或“Glyrichin 多肽”可互换使用，都指具有天然抗菌肽 Glyrichin 家族各成员氨基酸序列(序列表 1 中序列 1、序列 3-14)的蛋白或多肽。它们包括含有或不含起始甲硫氨酸的天然抗菌肽 Glyrichin 家族成员，以及
10 含有或不含有信号肽的 Glyrichin 成员蛋白。

术语“Glyrich 结构域”特指序列表中序列 1~13 中最为保守的一段多肽，其氨基酸长度在 59-68 之间，等电点大于 7，pH7.0 条件下带正电荷，Glycine 含量在 20 种氨基酸中含量为最高，并且含有至少一个疏水区的一段多肽。它们的序列及结构特点如图 3，图 4 和表
15 2 所界定。

术语“Glyrichin 家族”是指具有 Glyrich 结构域且在该结构域区段氨基酸同源性相互间在 30%以上的多肽，包括本发明中已经提到的 13 个多肽和其他尚未报道的多肽。以进化生物学的观点，它们来自
20 于同一个祖先。

本发明还提供了包含 Glyrichin 多肽或其片段的融合蛋白。除了几乎全长的多肽外，本发明还包括了 Glyrichin 多肽的可溶性片段。通常，该片段具有 Glyrichin 多肽序列中一定数目的连续氨基酸序列。应理解，本发明的多肽并不限于上述例举的代表性的多肽。

25 术语“修饰”(通常不改变一级结构)包括：体内或体外的多肽的化学形式如乙酰化或羧基化。修饰还包括糖基化，如那些在多肽的合成和加工或进一步加工步骤中进行糖基化修饰而产生的多肽。这种修饰可以通过将多肽暴露于进行糖基化的酶(如哺乳动物的糖基化酶

或去糖基化酶)而完成。修饰形式还包括具有磷酸化氨基酸残基(如磷酸丝氨酸、磷酸苏氨酸)的序列。还包括被修饰从而提高了其抗蛋白水解性能或优化了溶解性能的多肽。

在获得人源 Glyrichin 多核苷酸序列后,发明人以人源和鼠源 Glyrichin 为例,对该基因家族的功能开展研究,结果证明人源和鼠源 Glyrichin 是一个在人和小鼠多种组织中广泛表达的具有抗菌活性的天然免疫分子,是一种天然存在的抗菌肽,有望发展成为在医药工业及所有需要防治细菌感染不同领域的不同领域。这也是 Glyrichin 家族各个成员的基本功能和活性。

10

下面以人源 Glyrichin (人 Glyrichin, hGlyrichin) 基因的克隆、表达、纯化和抗菌活性测试实施例为例,进一步阐述本发明。应理解,这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法,通常按照常规条件如 Sambrook 等人的分子克隆:实验室手册(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件,或按照制造厂商所建议的条件。

实施例 1、人 Glyrichin 基因的获得及检测

1、人 Glyrichin 基因完整 ORF 序列的获得

根据鼠源 Glyrichin 基因设计如下一对简并引物并以人胎肝 RNA 为模板进行 RT-PCR 反应:

5' 引物 Pa: 5-CGATGCCGGTGGCCGTGGGTCCCT-3

3' 引物 Pb: 5-TTAGCATCGTATGCCCATTTCCA-3

PCR 反应体系: PCR 反应体系为分子生物学常规体系。PCR 扩增条件为: 94℃、4 分钟, 1 循环; 94℃、40 秒, 60℃、50 秒, 72℃、1 分钟, 30 循环; 72℃、7 分钟, 1 循环。PCR 产物经过 Winzard PCR preps purification kit (购自 Promega 公司) 纯化, 用 T₄ 连接酶重组到 pGEM-T 载体中得到重组质粒 pGEM-T/hGlyrichin, 然后转化大肠杆菌 JM109, 测序鉴定。结果表明人 Glyrichin 基因具有序列表中序列 2 的核苷酸序列,

编码序列具有表中序列 1 的氨基酸残基序列的富含甘氨酸蛋白人 Glyrichin。

2、Northern blot 分析

培养人 4 种肿瘤细胞株 HepG2、HeLa、Jurket 和 HEK293, 然后用 Winzard plus RNA purification kit (购自 Promega 公司) 提取总 RNA。各取 20 微克总 RNA, 在 1.2% 甲醛变性琼脂糖凝胶上分离, 并转到 Hybond-N+ 尼龙膜。hGlyrichin 完整 ORF 为探针, 用 Promega 公司的 Prime-a-gene 试剂盒标记。杂交结果如图 5A 所示, 表明 人 Glyrichin 基因在所测试的 4 种不同组织来源的肿瘤细胞株中都有表达, 提示它可能是一种广泛表达的天然抗菌肽。同时揭示只存在一个转录本, 且大小约为 600bp。

3、体外转录和翻译试验

以 pGEM-T/hGlyrichin 质粒为模板, 利用以下引物: 5' 引物: 5' -CGGGATCCCGATGCCGGTGGCCGTGGGTCCCT-3' 和 3' 引物: 5' -gctcgagtttagcatcggatgcccattcc-3' PCR 扩增 得到 hGlyrichin 基因全长的 ORF。其中 PCR 反应体系: 采用常规(参考精编分子生物学指南)PCR 反应体系。PCR 扩增条件为: 94℃、4 分钟, 1 循环; 94℃、40 秒, 60℃、50 秒, 72℃、1 分钟, 30 循环; 72℃、7 分钟, 1 循环。

然后用 BamHI 和 SacI 酶切 PCR 扩增产物; 将 pT7 载体(购自 Promega 公司)用同样的酶切割, 然后用 T₄ DNA 连接酶连接, 构建成 pT7-hGlyrichin 质粒。构建完成的质粒在如下体系中进行转录和翻译反应:

25 μl 体系:

兔网织红细胞裂解液: 12.5 μl

反应缓冲液: 1.0 μl

氨基酸混合物: 0.5 μl

25 S35-Met: 1.0 μl (50 μCi)

DNA 2.0 μl (0.5 μg/μl)

RNase 抑制剂: 0.5 μl

T7 DNA 聚合酶: 0.5 μl

去离子水: 7.0 μl

30 总体积: 25.0 μl

30℃ 反应 90 分钟。反应结束后取 5 μl 样品，加入 10 μl Loading Buffer，进行 SDS-PAGE 电泳，固定液中固定 30 分钟，干燥剂中浸泡 5 分钟，用干燥架固定凝胶，干燥凝胶过夜。放射自显影，-20℃压片 24 小时，洗片，结果如图 5B 所示，表明 hGlyrichin 基因编码的蛋白体外翻译后的大小约为 8.8kDa，同理论推断的结果一致，说明该基因可以在体外正常地转录和翻译。反应体系中所用材料来自 Promega 的 TNT 试剂盒，S35-Met 购自 Amersham Bioscience 公司。图中，1 为分子量标准，2 为无关基因 UBF (GenBankUBF-f1 AF294842；中国应用生理学杂志，20：66，2004) 的翻译产物，3 为 hGlyrichin 基因翻译产物，4 为试剂盒提供的阳性对照翻译产物。

实施例 2、内源性诱导表达的 hGlyrichin (mGlyrichin) 对大肠杆菌 BL21 的生长抑制试验

1、pET-22b-hGlyrichin 的构建

以 pGEM-T/hGlyrichin 为模板，以引物 1：5' -GGAATTCATATGCCGGT GGCCGTGGGTC-3' 和引物 2：5' -CCGCTCGAGTTAGCATCGGATGCCCATC-3' 常规 PCR 扩增 hGlyrichin 基因。其 PCR 反应体系（除引物外）和反应条件同实施例 1 的步骤 3。

将得到的 hGlyrichin 基因扩增产物用 Nde I 和 Xho I 酶切后，用 T₄ DNA 连接酶连入经 Nde I 和 Xho I 酶切的表达载体 pET-22b(+) 的 Nde I 和 Xho I 位点之间，转化大肠杆菌 BL21，通过酶切鉴定出阳性克隆 1 和 8（含有质粒 pET-22b-hGlyrichin）。

2、对照质粒 pET-22b (+) -UBF 和 pET-22b (+) -PTP 的构建

将 UBF 基因 (GenBankUBF-f1 AF294842；中国应用生理学杂志，20：66，2004) 和 PTP 基因 (甲状旁腺激素的 34 肽基因序列) (GenBank NM000315) 分别克隆入 pET-22b (+) 的多克隆位点得到对照质粒 pET-22b (+) -UBF 和 pET-22b (+) -PTP。

3、生长抑制试验

挑取阳性克隆，接种到 AMP 抗性的液体 LB 培养基（其中氨苄青霉素为 50ug/ml）中，于 37℃，250rpm 的摇床培养 12h；以 1：100 的体积比接入试管，继续培养至 OD 值为 0.03 时，向每只试管中加入 IPTG 至终浓

度为 0.5mM，阴性对照组则加入相应体积的 PBS；于 30℃、250rpm 摇菌，每隔 45 分钟取出 1 毫升菌液测 OD600 值，连续测 10 次以上。以时间为横轴，OD600 对数值为纵轴做生长曲线。结果如图 6A，图 6B，图 6C 和图 6D 所示，图 6A 表明单转空载体 pET-22b (+) 的菌株在加入和不加入 IPTG 两种情况下，细菌的生长曲线基本吻合；图 6B 表明在单转两个已知不具有抗菌活性基因 UBF 和 PTP 的情况下，加入和不加入 IPTG 两种情况下，细菌的生长同样不受影响；图 6C 和图 6D 表明，在转入 Glyrichin 后，加入和不加入 IPTG 两种情况下细菌的生长曲线明显不同，在前一种情况下细菌的生长明显受到抑制，从图中还可以看出，随着培养时间延长 IPTG 消耗殆尽，细菌的生长又得以恢复，而其他测试基因都没有此现象。上述结果充分说明这种抑菌效应是 Glyrichin 蛋白本身的作用。

实施例 3、游离 hGlyrichin 表达产物的纯化及杀菌活性检测

根据小鼠 Glyrichin 的 cDNA 序列设计引物 5' -CGGGATCCCGATGCCGGTGGCC GTGGGTCCCT-3' 和 5' -GGAATTCTTAGCATCGTATGCCCATTC-3'，以人胎肝 mRNA 进行 RT-PCR 扩增（其 PCR 反应体系（除引物外）和反应条件同实施例 1 的步骤 1），纯化后 PCR 产物经测序和限制性内切酶 BamHI 和 EcoRI 消化后在 T₄ DNA 连接酶的作用下，插入到用同样内切酶消化后的 pGEX4-4T2（购自 Pharmacia）原核表达载体中，经转化 JM109 大肠杆菌后，以 0.5mM IPTG 进行体外诱导表达 5 小时，收集菌体；用 PBS 重悬，反复冻融裂解；于 4℃、12000rpm 离心 20 分钟，取上清进行 SDS-PAGE 电泳。结果如图 7 所示，表明 GST-Glyrichin 融合蛋白在上清中表达，其分子量为 34KD。图中泳道 1 为蛋白质分子量标准，泳道 2 为 GST-hGlyrichin 融合蛋白诱导表达产物，泳道 3 为未经诱导 GST-hGlyrichin 融合蛋白表达产物，泳道 4 为 GST 蛋白诱导表达产物，泳道 5 为未诱导 GST 蛋白表达产物。

诱导表达得到的 GST-Glyrichin 融合蛋白，经 Sepharose 4B GST 纯化柱 (Pharmacia Inc.) 纯化后用肠激酶 (Enterokinase, Roche) 切割，得到游离的 hGlyrichin 蛋白，利用 96 孔板法检测 GST-Glyrichin 融合蛋白和游离的 hGlyrichin 蛋白的抗菌（大肠杆菌 DH5 α 和枯草杆菌 DB430）活性，测定最小抑菌浓度 (MIC)。在抗菌活性测定中，稀释菌液的浓度为 10⁴-10⁵CFU/ml，按每孔 80 微升菌液接种于 96 孔板中，将多肽以一定比例

稀释，每孔加入 5 微升；将 96 孔板置于 37℃ 培养 12h，紫外分析仪检测 OD600 值，结果显示，GST-Glyrichin 融合蛋白没有抑制细菌生长的活性，而游离的 Glyrichin 具有在较低浓度下抑制细菌生长的活性。游离的 Glyrichin 蛋白具有在较低浓度下抑制革兰氏阳性菌和革兰氏阴性菌的生长，具有抗菌活性（表 3）。

表 3. 游离 hGlyrichin 蛋白的最小抑菌浓度 (MIC)

| | 菌株名称 | 最小抑菌 浓度 (ug/ml) |
|--------|-------------------|--------------------|
| 革兰氏阴性菌 | 大肠杆菌 DH5 α | 0.2 |
| 革兰氏阳性菌 | 枯草杆菌 DB430 | 3 |

实施例 4、hGlyrichin 酵母表达及其表达产物杀菌活性的检测

通过 PCR 引物设计（全长上游引物：5' -AGGAATTCATGCCGGTGGCCGTGGG TCCCTAC-3'；5' 缺失体上游引物：5' -AGGAATTCATGGGCTTCGTGATGGGTTGC -3'；全长下游引物：5' -AAGGAAAAAAGCGGCCGCTTAGCATCGGATGCCCATCCCA ATG-3'），以含 hGlyrichin 全长的 pET-22b 质粒为模板，在全长上游引物和全长下游引物、5' 缺失体上游引物和全长下游引物的引导下分别 PCR 扩增全长 hGlyrichin 基因和 5' 端缺失序列 2 的自 5' 端第 1 到第 60 碱基的 5' 缺失体基因。其 PCR 反应体系（除引物外）和反应条件同实施例 1 的步骤 3。

将 PCR 扩增获得的全长 hGlyrichin 基因和 5' -缺失体基因分别插入到 pPIC9K 酵母表达质粒（Invitrogen）的 EcoRI 和 NotI 之间，经转化大肠杆菌 BL-21 工程菌和克隆筛选，获得构建后分别带有全长和 5' 缺失的 hGlyrichin 基因的质粒。所获得的高纯度质粒经 SalI 内切酶线性化后，采用电转法转化 GS115 酵母菌，经含 G418(50ug/ml) 的 MD 平板筛选获得阳性克隆。将阳性克隆接种到含 5ml BMGY 培养基中，30℃ 摇床培养至 OD600=2.0-6.0 时，然后用 BMMY（含终浓度为 1% 的甲醇）培养基稀释到 OD600=1.0，继续培养。以后每 24 小时补加甲醇到终浓度为 0.5%，诱导培养后不同时间取 1ml 加入到 Eppendorff

离心管中，离心 16000 转，留上清作活性测试。活性测试采用琼脂平板扩散法，即在含有 1%琼脂的 LB 平板表面均匀涂布测试菌（大肠杆菌 BL-21），再在每个小方块中央放置直径约 2mm 的滤纸片，将 15ul 不同克隆上清滴加到纸片上，20 分钟后等量加入，共加 3 次。37℃培养 4 小时观察抑菌环的大小以获得阳性克隆。同时以 10ul 浓度为 100mg/ml 的氨苄青霉素作为活性对照。经过活性筛选，获得了表达全长 hGlyrichin 的阳性克隆（L4 克隆）和表达 5' -缺失 hGlyrichin 的阳性克隆（S2, S12 和 S5 克隆）。L4, S2, S12 和 S5 克隆的抑菌效果如图 8 所示，表明转化了无论含全长还是 5' -缺失的 hGlyrichin 基因的阳性克隆（L4, S2, S12 和 S5）上清具有明显抑制琼脂板表面大肠杆菌 BL-21 生长的效果。上述实验结果证实，酵母表达系统可以表达 hGlyrichin 蛋白，而且，表达产物具有抗菌活性；全长和 5' -缺失的基因均表达活性蛋白，为进一步寻找代表该基因活性的确切结构奠定了牢固的基础。根据上述系列实验结果，业内人员极易利用类似方法进一步筛选获得具有相似活性的最小蛋白质分子。

实施例 5、hGlyrichin 基因的构-效关系研究

1、不同缺失体质粒转化受体菌对受体菌生长的影响

发明人利用基因重组技术，设计了系列 PCR 引物：

引物 1：5' -GGAATTCATATGCCGGTGGCCGTGGGTC-3'

引物 2：5' -GGAATTCATATGGGCTTCGTGATGGGTTGC-3'

引物 3：5' -CCGGCTC GAG TTA GAA TGT GCC AAA GGT-3'

引物 4：5' -CCGCTCGAGTTAGCATCGGATGCCCATC-3'

以 pET-22b-hGlyrichin 为模板，在不同引物组合的引导下分别 PCR 扩增全长 hGlyrichin 基因（引物 1 和引物 4 组合）、5' 端缺失序列 2 的自 5' 端第 1-60 位碱基的 5' 缺失体基因（引物 2 和引物 4 组合）和 3' 端缺失序列 2 的自 5' 端第 211-240 位碱基的 3' 缺失体基因（引物 1 和引物 3 组合）。其 PCR 反应体系（除引物外）和反应

条件同实施例 1 的步骤 3。将 PCR 扩增获得的不同长度 hGlyrichin 基因片段，按照实施例 2 中叙述的方法分别进行基因片段与 pET-22b(+) 质粒连接、转化大肠杆菌 BL-21 菌、挑克隆并鉴定。同样，按照实验例 2 方法进行上述阳性克隆菌的诱导表达以观察全长及不同缺失体对转化菌生长的影响。结果如图 9 所示，表明转化了空载体的大肠杆菌 BL-21 无论加或不加 IPTG，其生长不受影响；而转化了带全长或部分缺失的目的基因的质粒后，不加 IPTG 诱导，细菌生长不受影响，IPTG 诱导后 5 小时内，转化菌的生长明显被抑制。图中 X 轴为 IPTG 诱导后小时数(括号中)，Y 轴为 OD600 光吸收值。上述结果证实，hGlyrichin 基因的 5' -端和 3' -端部分核苷酸序列对于发挥抗菌作用不是必需的。图 9 中各个图例中的 (—) 表示未加 IPTG，(+) 表示加 0.5mM IPTG。

工业应用

发明人通过实验，首次发现了 Glyrichin 家族，得到了其保守序列，并证明了 hGlyrichin 具有抗细菌的作用。Glyrichin 可以应用于医药领域和所有需要应用抗生素的不同领域，有广阔的应用前景。

权利要求

1、富含甘氨酸蛋白，选自下述蛋白质家族中的至少一种：

5 1) 人 Glyrichin 和小鼠 Glyrichin：均具有序列表中序列 1 的氨基酸残基序列的蛋白质或将序列表中序列 1 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

10 2) 斑马鱼 (*Danio rerio*) Glyrichin：具有序列表中序列 3 的氨基酸残基序列的蛋白质或将序列表中序列 3 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

15 3) 按蚊 (*Anopheles gambiae*) Glyrichin：具有序列表中序列 4 的氨基酸残基序列的蛋白质或将序列表中序列 4 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

4) 果蝇 (*Drosophila melanogaster*) Glyrichin：具有序列表中序列 5 的氨基酸残基序列的蛋白质或将序列表中序列 5 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

20 5) 线虫 (*Caenorhabditis elegans*) Glyrichin：具有序列表中序列 6 的氨基酸残基序列的蛋白质或将序列表中序列 6 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

25 6) 线虫 (*Caenorhabditis elegans*) Glyrichin：具有序列表中序列 7 的氨基酸残基序列的蛋白质或将序列表中序列 7 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

30 7) 芽殖酵母 (*Schizosaccharomyces pombe*) Glyrichin：具有序列表中序列 8 的氨基酸残基序列的蛋白质或将序列表中序列 8 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

8) 酿酒酵母 (*Saccharomyces cerevisiae*) Glyrichin：具有序列表中

序列 9 的氨基酸残基序列的蛋白质或将序列表中序列 9 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

9) 拟南芥 (*Arabidopsis thaliana*) Glyrichin: 具有序列表中序列 10 的氨基酸残基序列的蛋白质或将序列表中序列 10 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

10) 疟原虫 (*Plasmodium falciparum 3D7*) Glyrichin: 具有序列表中序列 11 的氨基酸残基序列的蛋白质或将序列表中序列 11 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

11) 疟原虫 (*Plasmodium yoelii yoelii*) Glyrichin: 具有序列表中序列 12 的氨基酸残基序列的蛋白质或将序列表中序列 12 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

12) 稻瘟病菌 (*Magnaporthe grisea*) Glyrichin: 具有序列表中序列 13 的氨基酸残基序列的蛋白质或将序列表中序列 13 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质；

13) 脉孢菌 (*Neurospora crassa*) Glyrichin: 具有序列表中序列 14 的氨基酸残基序列的蛋白质或将序列表中序列 14 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质。

2、根据权利要求 1 所述的富含甘氨酸蛋白，其特征在于：所述富含甘氨酸蛋白为人 Glyrichin 和小鼠 Glyrichin，它为具有序列表中序列 1 的氨基酸残基序列的蛋白质或将序列表中序列 1 的氨基酸残基序列经过 1 至 20 个氨基酸残基的缺失、插入和/或取代以及在羧基末端和/或氨基末端添加 1 至 20 个氨基酸残基且具有抗细菌作用的蛋白质。

3、根据权利要求 1 或 2 所述的富含甘氨酸蛋白，其特征在于：所述缺失、插入和/或取代以及在羧基末端和/或氨基末端添加的氨基酸残基的数目为 1 至 10 个。

4、根据权利要求 3 所述的富含甘氨酸蛋白，其特征在于：所述缺失、

插入和/或取代以及在羧基末端和/或氨基末端添加的氨基酸残基的数目为 1 至 5 个。

- 5 5、根据权利要求 4 所述的富含甘氨酸蛋白，其特征在于：所述缺失、插入和/或取代以及在羧基末端和/或氨基末端添加的氨基酸残基的数目为 1 至 3 个。

6、权利要求1-5任一所述的富含甘氨酸蛋白的编码基因。

- 7、根据权利要求 6 所述的基因，其特征在于：所述富含甘氨酸蛋白为人 Glyrichin，其编码基因具有序列表中 SEQ ID No: 2 的 DNA 序列或与序列表中 SEQ ID No: 2 限定的 DNA 序列具有 90% 以上同源性，且编码序列表中序列 1 的氨基酸残基序列的 DNA 序列或在高严谨条件下可与序列表中的序列 2 限定的 DNA 序列杂交的核苷酸序列。

8、含有权利要求6或7所述基因的表达载体。

9、含有权利要求6或7所述基因的细胞系。

10、含有权利要求6或7所述基因的工程菌。

- 15 11、权利要求1-5任一所述的富含甘氨酸蛋白及其编码基因在抗细菌中的应用。

12、根据权利要求11所述的应用，其特征在于：所述富含甘氨酸蛋白及其编码基因用于制备预防和/或治疗人或畜细菌性感染疾病的药物。

- 20 13、根据权利要求 11 所述的应用，其特征在于：所述富含甘氨酸蛋白及其编码基因用于制备具有预防和/或治疗潜在细菌感染的不同种类生物用制品。

14、根据权利要求 11 所述的应用，其特征在于：所述富含甘氨酸蛋白及其编码基因用于生产抗病虫害的转基因生物。

- 25 15、根据权利要求 11 所述的应用，其特征在于：所述富含甘氨酸蛋白及其编码基因用于制备针对所述富含甘氨酸蛋白的衍生物或拮抗剂及其配体、抗体。

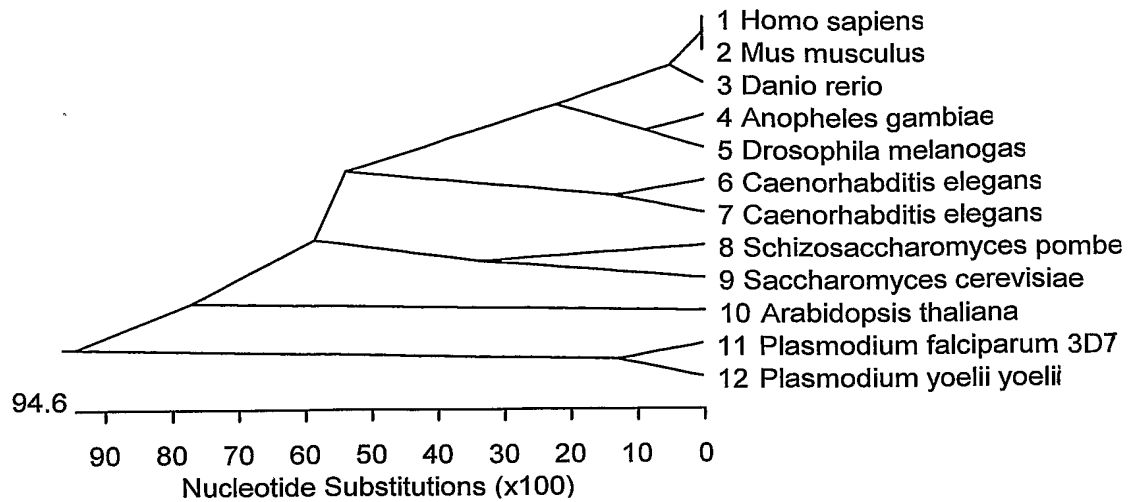


图 1

| Majority | QPSC | FD | KVKMGFMGFAVGGATGALFGGFSALRIGMRGRELMLGIG | KTMQSGGTGTFGTMGIGMGIRC | | | | | |
|-------------------|--|-----------------------|---|------------------------|-----------------------|----|----|----|--|
| | | 10 | 20 | 30 | 40 | 50 | 60 | 70 | |
| H. sapiens | QPSC | FD | RVKMGFMGFAVGGATGALFGGFSALRIGMRGRELMLGIG | KTMQSGGTGTFGTMGIGMGIRC | 68 | | | | |
| M. musculus | QPSC | FD | RVKMGFMGFAVGGATGALFGGFSALRIGMRGRELMLGIG | KTMQSGGTGTFGTMGIGMGIRC | 68 | | | | |
| D. rerio | QPSC | FD | RVKMGFMGFAVGGATGALFGGFSALRIGMRGRELMLGIG | KTMQSGGTGTFGTMGIGMGIRC | 68 | | | | |
| A. gambiae | QPSC | FD | RMKMGFMGFAVGGATGALFGGFSALRIGMRGRELMLGIG | KTMQSGGTGTFGTMGIGMGIRC | 68 | | | | |
| D. melanogaster | GPSC | FD | RMKMGFMGFAVGGATGALFGGFSALRIGMRGRELMLGIG | KTMQSGGTGTFGTMGIGMGIRC | 68 | | | | |
| C. elegans | DKICRNLTLYFQNFQIRGCLMGGAMIGCATGHLGGTMGFRAGMRCKDELQTC | KVYAGCGSPGVFMGVAQCLRY | 75 | | | | | | |
| C. elegans2 | QPSC | FD | RVKMGFMGFAVGGATGALFGGFSALRIGMRGRELMLGIG | KTMQSGGTGTFGTMGIGMGIRC | 68 | | | | |
| A. thaliana | KNBC | LA | KITAQVANGGELGGVGVVYTYEIRIVKVPDLHKVRFIC | QHTSSAAIDCLDCAESLHC | 68 | | | | |
| S. cerevisiae | QFEN | MD | KFMGELMGGTINVCYSIRGGRATATCGPGDGVVRTIC | KYIAGGATLLEDSISILIS | 68 | | | | |
| S. pombe | QEST | VE | KLHSAIDGSALEIGTFFHFGVAVIVYEPGPGFLRTIC | QYMLTSAANGFFHSISVLIIN | 68 | | | | |
| M. grisea | ABEN | MD | KPVCATMGFIYCTVIVFRGGAGPN | GIRRTIC | QYMLTSAANGFFHSISVLIIN | 59 | | | |
| P. falciparum 3D7 | TRAC | FE | SITMGVKGIMGGIFGSLTGLYASF | AHKMFHLEPVSIVGGAVS | GGFELCCGIVRC | 63 | | | |
| P. yoelii | TKAC | LE | SVKMGVKGIMGGIFGSLTGLYASF | AHKMFHLEPVSIVGGAVS | GGFELCCGIVRC | 63 | | | |

图 2

| | | Percent Identity | | | | | | | | | | | | | | | Divergence | |
|----|-------|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|----|--|---------------------------|------------|--|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | | | | |
| 1 | | 100.0 | 94.7 | 76.3 | 74.1 | 55.3 | 42.1 | 50.0 | 51.3 | 40.8 | 35.5 | 35.5 | 42.1 | 1 | | Homo sapiens | | |
| 2 | 0.0 | | 94.7 | 76.3 | 74.1 | 55.3 | 42.1 | 50.0 | 51.3 | 40.8 | 35.5 | 35.5 | 42.1 | 2 | | Mus musculus | | |
| 3 | 6.1 | 6.1 | | 77.6 | 73.7 | 55.3 | 42.1 | 50.0 | 60.0 | 39.5 | 32.9 | 32.9 | 42.1 | 3 | | Danio rerio | | |
| 4 | 32.7 | 32.7 | 30.4 | | 90.8 | 51.3 | 38.2 | 48.7 | 60.0 | 35.5 | 32.9 | 31.6 | 38.2 | 4 | | Anopheles gambiae | | |
| 5 | 42.2 | 42.2 | 37.3 | 11.1 | | 48.7 | 36.8 | 44.7 | 47.4 | 38.2 | 30.3 | 28.9 | 39.5 | 5 | | Drosophila melanogaster | | |
| 6 | 79.9 | 79.9 | 79.9 | 92.5 | 102.0 | | 84.2 | 48.7 | 40.8 | 36.8 | 38.2 | 35.5 | 36.8 | 6 | | Caenorhabditis elegans | | |
| 7 | 92.5 | 92.5 | 92.5 | 107.1 | 112.6 | 7.8 | | 34.2 | 26.3 | 25.0 | 28.9 | 27.6 | 23.7 | 7 | | Caenorhabditis elegans | | |
| 8 | 97.1 | 97.1 | 97.1 | 102.0 | 118.4 | 102.0 | 124.6 | | 63.2 | 35.5 | 35.5 | 35.5 | 44.7 | 8 | | Saccharomyces cerevisiae | | |
| 9 | 92.5 | 92.5 | 97.1 | 97.1 | 107.1 | 138.4 | 174.0 | 69.0 | | 39.5 | 32.9 | 31.6 | 51.3 | 9 | | Schizosaccharomyces pombe | | |
| 10 | 138.4 | 138.4 | 146.1 | 174.0 | 154.5 | 163.8 | 185.4 | 174.0 | 146.1 | | 39.5 | 38.2 | 34.2 | 10 | | Arabidopsis thaliana | | |
| 11 | 146.6 | 146.6 | 166.3 | 166.3 | 191.1 | 130.3 | 130.3 | 146.6 | 166.3 | 123.2 | | 92.1 | 30.3 | 11 | | Plasmodium falciparum 3D7 | | |
| 12 | 146.6 | 146.6 | 166.3 | 177.9 | 202.0 | 146.6 | 138.4 | 146.6 | 177.9 | 130.3 | 10.2 | | 28.9 | 12 | | Plasmodium yoelli yoelli | | |
| 13 | 108.9 | 108.9 | 108.9 | 124.8 | 122.3 | 138.0 | 167.8 | 97.3 | 73.3 | 156.8 | 194.2 | 207.0 | | 13 | | Magnaporthe oryzae | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | | | | | |

图 3

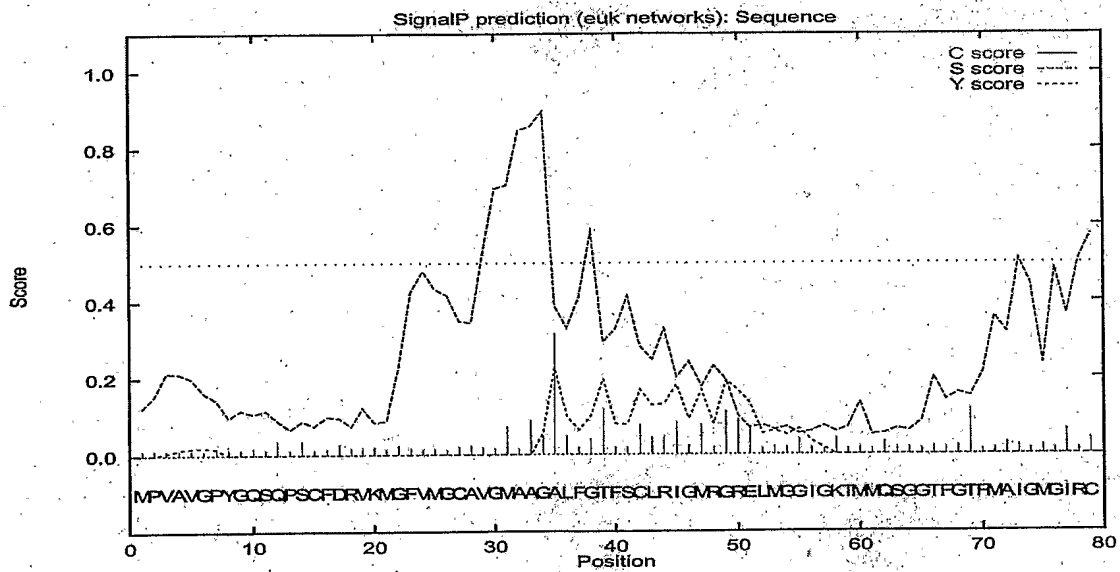


图 4

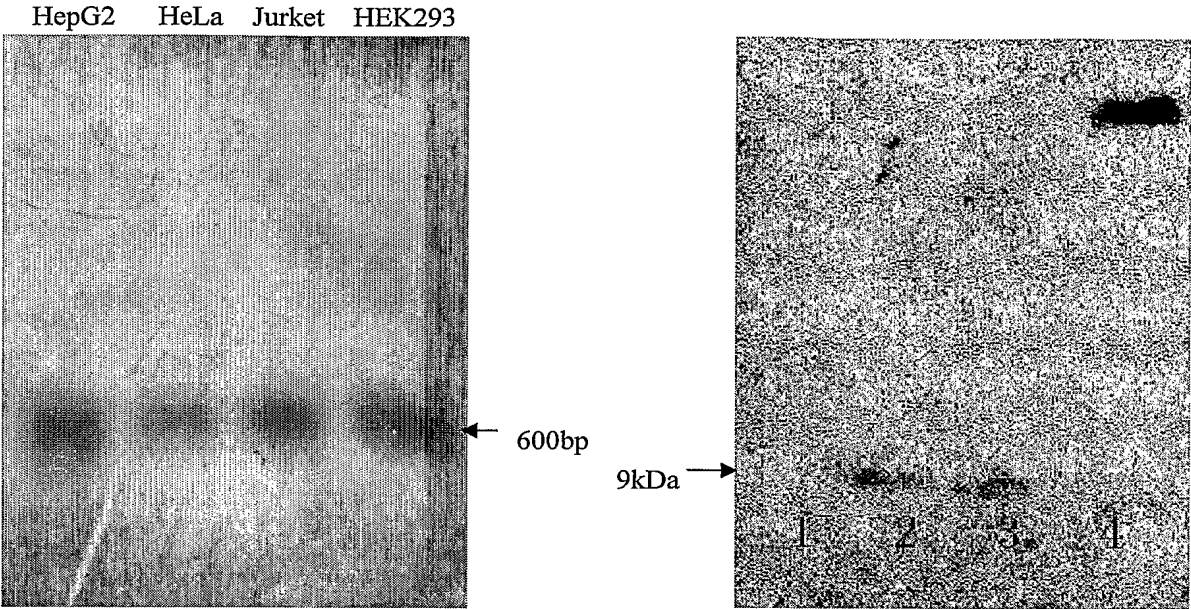


图 5A

图 5B

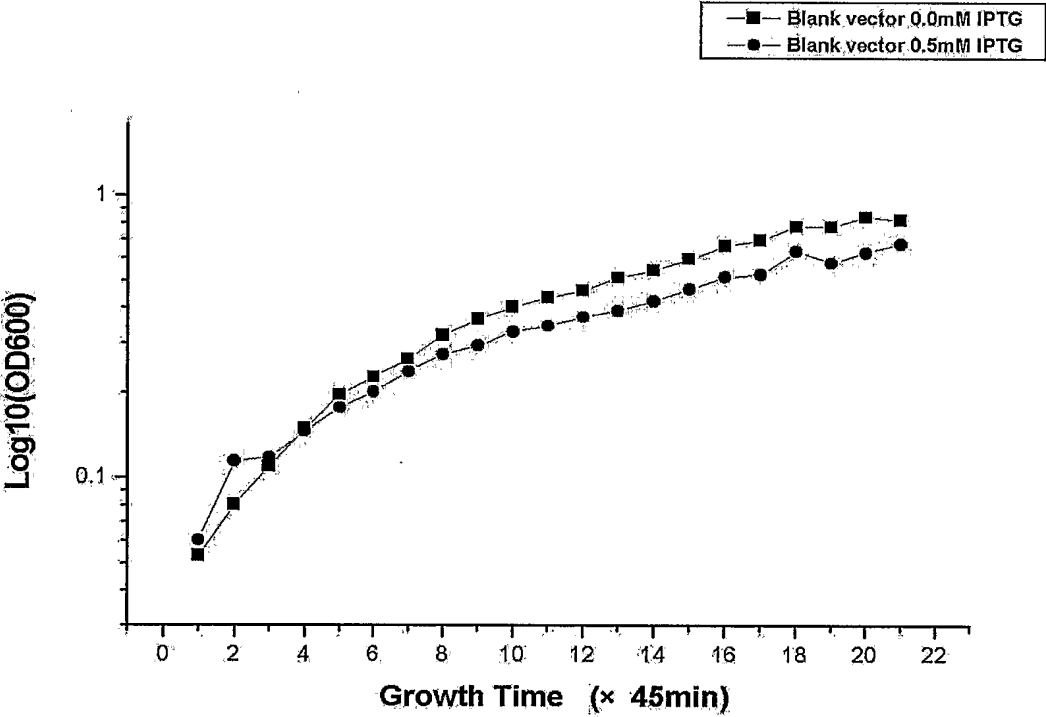


图 6A

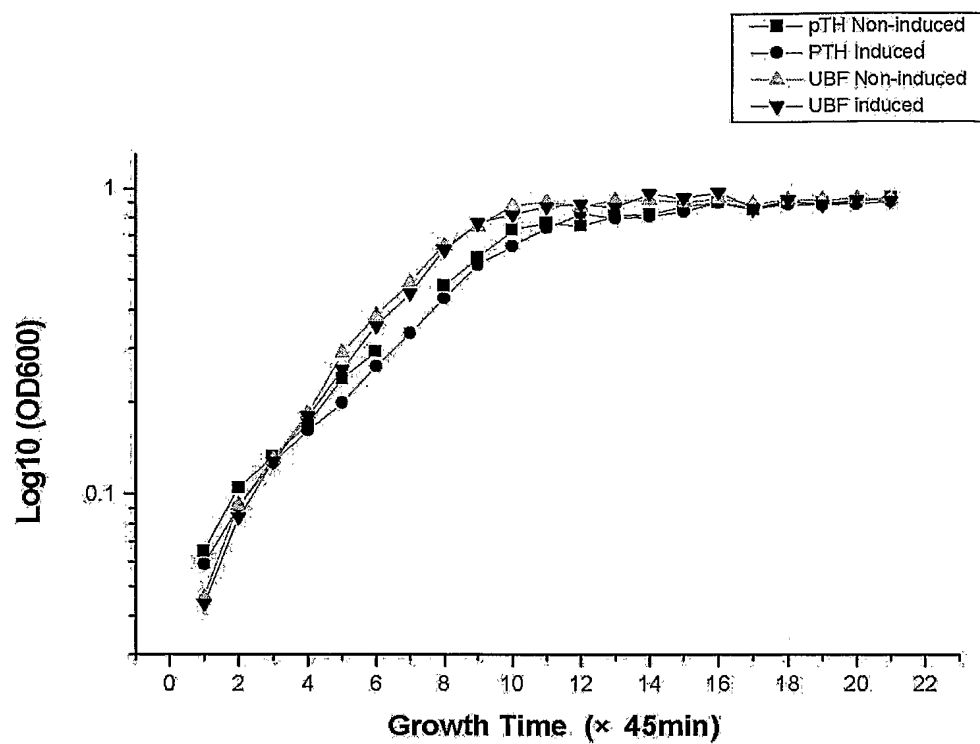


图 6B

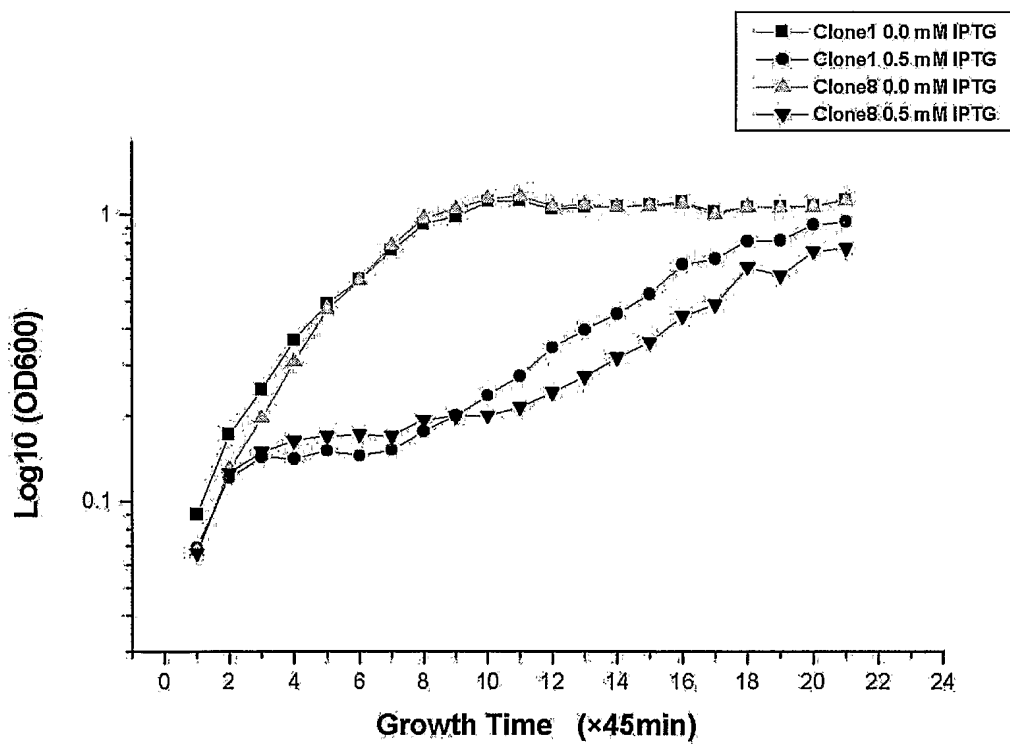


图 6C

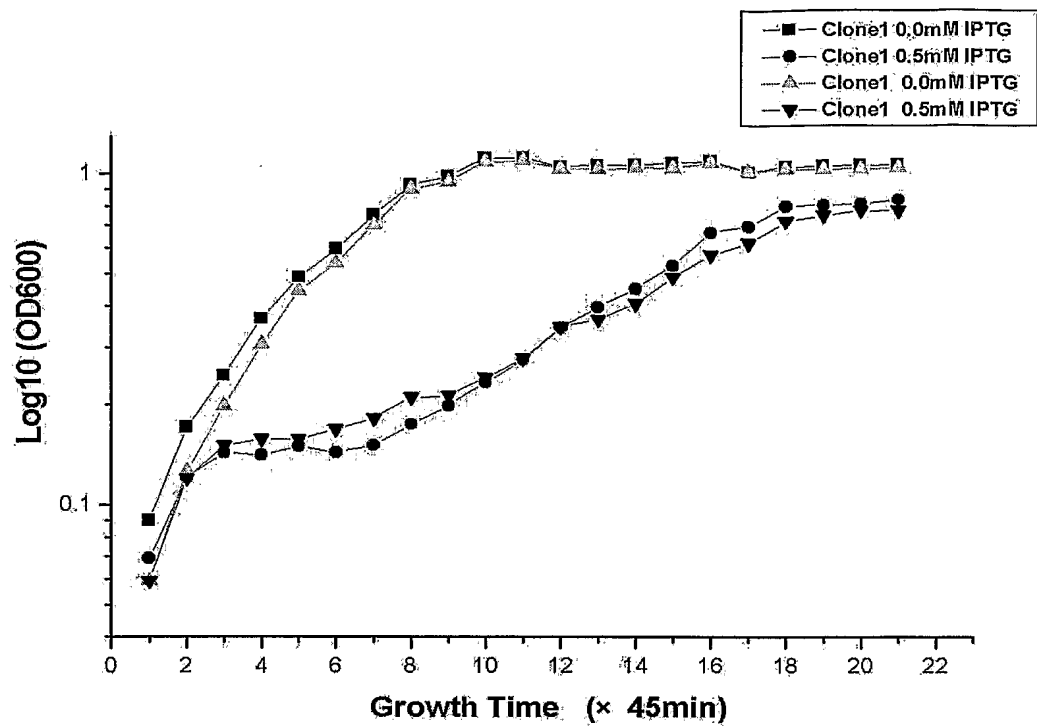


图 6D

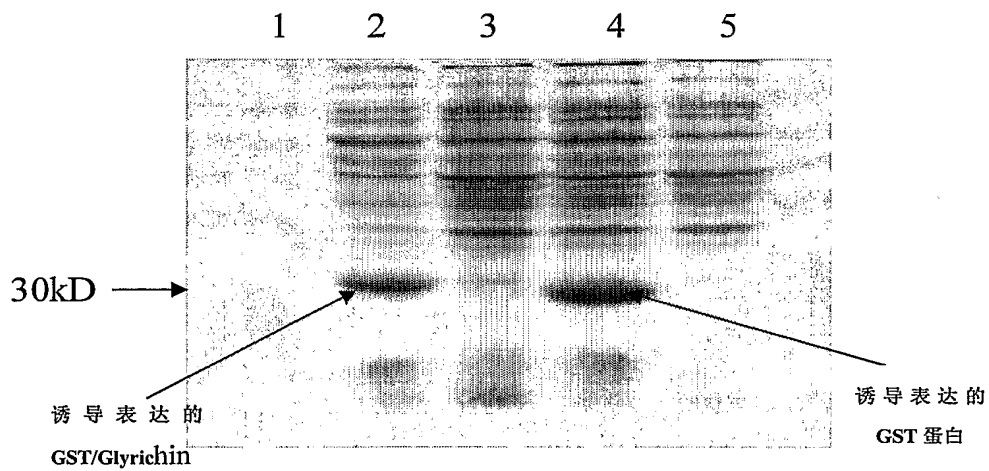


图 7

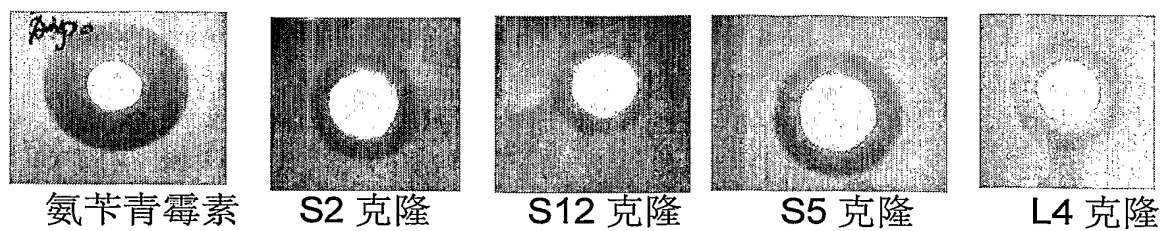


图 8

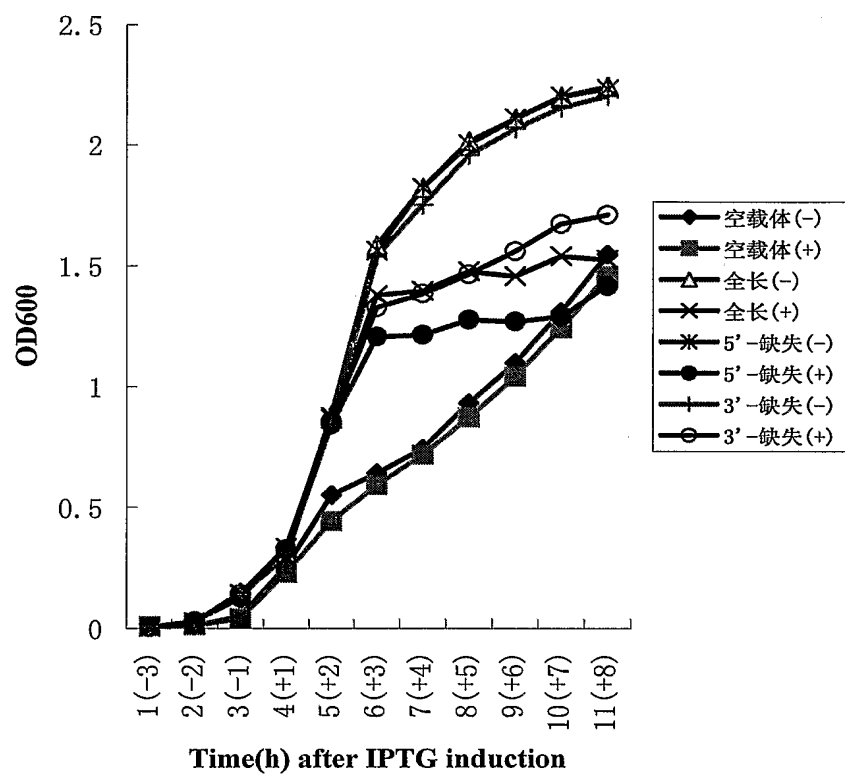


图 9

序列表

5 <160> 14
 <210> 1
 <211> 79
 <212> PRT
 <213> 人属人 (*Homo sapiens*)
 10 <400> 1
 Met Pro Val Ala Val Gly Pro Tyr Gly Gln Ser Gln Pro Ser Cys Phe
 1 5 10 15
 Asp Arg Val Lys Met Gly Phe Val Met Gly Cys Ala Val Gly Met Ala
 15 20 25 30
 Ala Gly Ala Leu Phe Gly Thr Phe Ser Cys Leu Arg Ile Gly Met Arg
 35 40 45
 Gly Arg Glu Leu Met Gly Gly Ile Gly Lys Thr Met Met Gln Ser Gly
 50 55 60
 20 Gly Thr Phe Gly Thr Phe Met Ala Ile Gly Met Gly Ile Arg Cys
 65 70 75
 <210> 2
 <211> 240
 25 <212> DNA
 <213> 人属人 (*Homo sapiens*)
 <400> 2
 atgccggtgg ccgtgggtcc ctacggacag tcccagccaa gctgcttcga ccgtgtcaaa 60
 30 atgggcttcg tgatgggttg cgccgtgggc atggcgccg gggcgctctt cggcaccttt 120
 tcctgtctca ggatcggaat gcggggtcga gagctgatgg gcggcattgg gaaaaccatg 180
 atgcagagtg gcggcacctt tggcacattc atggccattg ggatgggcat ccgatgctaa 240
 35 <210> 3

<211> 80

<212> PRT

<213> 斑马鱼 (*Danio rerio*)

5 <400> 3

Met Pro Val Ser Val Gly Ser Tyr Gly Gln Gln Ala Gln Pro Ser Cys
 1 5 10 15
 Phe Asp Arg Val Lys Met Gly Phe Met Met Gly Phe Ala Val Gly Met
 20 25 30
 10 Ala Ala Gly Ala Met Phe Gly Thr Phe Ser Cys Leu Arg Ile Gly Met
 35 40 45
 Arg Gly Arg Glu Leu Met Gly Gly Val Gly Lys Thr Met Met Gln Ser
 50 55 60
 Gly Gly Thr Phe Gly Thr Phe Met Ala Ile Gly Met Gly Ile Arg Cys
 15 65 70 75 80

<210> 4

<211> 128

<212> PRT

20 <213> 按蚊 (*Anopheles gambiae*)

<400> 4

Tyr Tyr Tyr Val Ile Val Val His Cys Cys Asp Asn Thr His Phe Asn
 1 5 10 15
 25 Glu Phe Val Pro Lys Ile Lys Leu Pro Pro Arg Lys Arg Tyr Val Arg
 20 25 30
 Ser Gly Ser Phe Gln Ile Leu Gln Lys Thr Asp Thr Lys Ser Thr Met
 35 40 45
 Pro Ala Val Pro Gly Gly Val Tyr Ser Gln Asn Gln Gln Pro Ser Cys
 30 50 55 60
 Phe Asp Arg Met Lys Met Gly Phe Thr Ile Gly Phe Cys Val Gly Met
 65 70 75 80
 Ala Ser Gly Ala Leu Phe Gly Gly Phe Ser Ala Leu Arg Tyr Gly Leu
 85 90 95
 35 Arg Gly Arg Glu Leu Ile Asn Asn Val Gly Lys Val Met Val Gln Gly

100 105 110
 Gly Gly Thr Phe Gly Thr Phe Met Ala Ile Gly Thr Gly Ile Arg Cys
 115 120 125

5 <210> 5
 <211> 79
 <212> PRT
 <213> 果蝇 (*Drosophila melanogaster*)

10 <400> 5
 Met Pro Leu Pro Thr Ser Ser Phe Ser Gln Gln Gly Pro Thr Cys Phe
 1 5 10 15
 Asp Lys Met Lys Thr Gly Phe Ile Ile Gly Phe Cys Val Gly Met Ala
 20 25 30
 15 Ser Gly Ala Val Phe Gly Gly Phe Ser Ala Leu Arg Tyr Gly Leu Arg
 35 40 45
 Gly Arg Glu Leu Ile Asn Asn Val Gly Lys Thr Met Val Gln Gly Gly
 50 55 60
 Gly Thr Phe Gly Thr Phe Met Ala Ile Gly Thr Gly Ile Arg Cys
 20 65 70 75

<210> 6
 <211> 145
 <212> PRT
 25 <213> 线虫 (*Caenorhabditis elegans*)

<400> 6
 Met Pro Val Pro Ser Gly Tyr Ala Ala His Pro Gln Gly Ser Gln Pro
 1 5 10 15
 30 Ser Cys Phe Thr Lys Ile Arg Met Gly Leu Met Met Gly Ala Met Ile
 20 25 30
 Gly Gly Ala Thr Gly Ile Leu Leu Gly Gly Phe Met Gly Phe Arg Ala
 35 40 45
 Gly Met Arg Gly Lys Asp Leu Leu Leu Gln Thr Gly Lys Thr Val Ala
 35 50 55 60

Gln Ser Gly Gly Ser Phe Gly Val Phe Met Gly Val Ala Gln Gly Leu
 65 70 75 80
 Arg Tyr Ile Phe Phe Lys Asn Leu Ala Gly Thr Gly Phe Trp Pro Phe
 85 90 95
 5 Ser Leu Asn Phe Ser Arg Ser Ile Asp Asn Cys Pro Arg Gly Lys Val
 100 105 110
 Val Tyr Ser Thr Arg Thr Asn Ala Phe Arg Phe Thr Thr Glu Ile Glu
 115 120 125
 Lys Lys Glu Pro Arg Arg Asp Thr Gln Arg Ala Val Asn Leu Pro Gln
 10 130 135 140
 Ile
 145

<210> 7
 15 <211> 162
 <212> PRT
 <213> 线虫 (*Caenorhabditis elegans*)

<400> 7
 20 Met Gln His Thr His Lys Glu Ala Asn Arg Arg Val Leu Gln Arg Lys
 1 5 10 15
 Lys Ile Asn Leu Leu Glu Met Ser Asp Lys Ile Cys Arg Asn Leu Ile
 20 25 30
 Tyr Phe Gln Asn Phe Gln Ile Arg Met Gly Leu Met Met Gly Ala Met
 25 35 40 45
 Ile Gly Gly Ala Thr Gly Ile Leu Leu Gly Gly Phe Met Gly Phe Arg
 50 55 60
 Ala Gly Met Arg Gly Lys Asp Leu Leu Leu Gln Thr Gly Lys Thr Val
 65 70 75 80
 30 Ala Gln Ser Gly Gly Ser Phe Gly Val Phe Met Gly Val Ala Gln Gly
 85 90 95
 Leu Arg Tyr Ile Phe Phe Lys Asn Leu Ala Gly Thr Gly Phe Trp Pro
 100 105 110
 Phe Ser Leu Asn Phe Ser Arg Ser Ile Asp Asn Cys Pro Arg Gly Lys
 35 115 120 125

Val Val Tyr Ser Thr Arg Thr Asn Ala Phe Arg Phe Thr Thr Glu Ile
 130 135 140
 Glu Lys Lys Glu Pro Arg Arg Asp Thr Gln Arg Ala Val Asn Leu Pro
 145 150 155 160
 5 Gln Ile

<210> 8

<211> 120

<212> PRT

10 <213> 芽殖酵母 (*Schizosaccharomyces pombe*)

<400> 8

Met Gln Ser Met Gln Pro Ser Thr Val Asp Lys Leu Lys Met Gly Ala
 1 5 10 15
 15 Ile Met Gly Ser Ala Ala Gly Leu Gly Ile Gly Phe Leu Phe Gly Gly
 20 25 30
 Val Ala Val Leu Arg Tyr Gly Pro Gly Pro Arg Gly Phe Leu Arg Thr
 35 40 45
 Leu Gly Gln Tyr Met Leu Thr Ser Ala Ala Thr Phe Gly Phe Phe Met
 20 50 55 60
 Ser Ile Gly Ser Val Ile Arg Asn Glu Asp Ile Pro Leu Ile Gln Gln
 65 70 75 80
 Ser Gly Ser His Trp Asn Gln Arg Leu Leu Asn Glu Asn Ala Asn Ser
 85 90 95
 25 Ser Arg Ile Phe Ala Leu Ala Met Gln Gln Ala Lys Ser Ser Pro Arg
 100 105 110
 Lys Ser Asn Glu Val Ala Glu Cys
 115 120

30 <210> 9

<211> 113

<212> PRT

<213> 酿酒酵母 (*Sacchchromyces cerevisiae*)

35 <400> 9

Met Pro Pro Leu Pro Gln Asn Tyr Ala Gln Gln Gln Pro Ser Asn Trp
 1 5 10 15
 Asp Lys Phe Lys Met Gly Leu Met Met Gly Thr Thr Val Gly Val Cys
 20 25 30
 5 Thr Gly Ile Leu Phe Gly Gly Phe Ala Ile Ala Thr Gln Gly Pro Gly
 35 40 45
 Pro Asp Gly Val Val Arg Thr Leu Gly Lys Tyr Ile Ala Gly Ser Ala
 50 55 60
 Gly Thr Phe Gly Leu Phe Met Ser Ile Gly Ser Ile Ile Arg Ser Asp
 10 65 70 75 80
 Ser Glu Ser Ser Pro Met Ser His Pro Asn Leu Asn Leu Gln Gln Gln
 85 90 95
 Ala Arg Leu Glu Met Trp Lys Leu Arg Ala Lys Tyr Gly Ile Arg Lys
 100 105 110
 15 Asp

<210> 10

<211> 74

<212> PRT

20 <213> 拟南芥 (*Arabidopsis thaliana*)

<400> 10

Met Ala Lys Asn Ser Cys Leu Ala Lys Ile Thr Ala Gly Val Ala Val
 1 5 10 15
 25 Gly Gly Ala Leu Gly Gly Ala Val Gly Ala Val Tyr Gly Thr Tyr Glu
 20 25 30
 Ala Ile Arg Val Lys Val Pro Gly Leu His Lys Val Arg Phe Ile Gly
 35 40 45
 Gln Thr Thr Leu Ser Ser Ala Ala Ile Phe Gly Leu Phe Leu Gly Ala
 30 50 55 60
 Gly Ser Leu Ile His Cys Gly Lys Gly Tyr
 65 70

<210> 11

35 <211> 168

<212> PRT

<213> 疟原虫 (*Plasmodium falciparum* 3D7)

<400> 11

5 Met Met Asn Trp Phe Arg Lys Lys Asp Ser Asn Glu Asn Lys Lys Val
 1 5 10 15
 Lys Ser Glu Tyr Asp Glu Tyr Val Thr Pro Pro Pro Phe Gly Asn Tyr
 20 25 30
 Leu Val Ser Glu Pro Lys Lys Pro Lys Ser Leu Lys Asn Asp Lys Thr
 10 35 40 45
 Ala Ile Thr Glu Phe Lys Gly Phe Thr Pro Pro Pro Lys Phe Glu Phe
 50 55 60
 Lys Glu Asp Ile Ser Asp Asn Lys Tyr Glu Glu Asp Phe Ser Lys Tyr
 65 70 75 80
 15 Thr Ser Asn Asn Ile Ile Asp Ser Ser Phe Tyr Asp Asp Lys Lys Lys
 85 90 95
 Leu Ser Asp Val Asn Leu Ser His Arg Thr Arg Ala Cys Phe Glu Ser
 100 105 110
 Ile Lys Met Gly Val Lys Met Gly Thr Met Val Gly Gly Ile Phe Gly
 20 115 120 125
 Ser Leu Thr Gly Ile Tyr Ala Ser Phe Ala His Lys Asn Leu Phe Ile
 130 135 140
 Leu Pro Val Ser Val Leu Gly Gly Ala Val Ser Phe Gly Phe Phe Leu
 145 150 155 160
 25 Gly Cys Gly Met Ile Val Arg Cys
 165

<210> 12

<211> 167

30 <212> PRT

<213> 疟原虫 (*Plasmodium yoelii yoelii*)

<400> 12

Met Met Asn Trp Phe Lys Lys Lys Glu Thr Thr Glu Glu Pro Gln Val
 35 1 5 10 15

Lys Ser Glu Tyr Asp Ser Tyr Val Thr Pro Pro Pro Phe Gly Asn Tyr
 20 25 30
 Leu Ala Lys Lys Pro Glu Lys Pro Lys Ser Leu Lys Asn Glu Lys Ile
 35 40 45
 5 Asn Val Thr Glu Phe Lys Gly Phe Thr Pro Pro Pro Lys Phe Glu Phe
 50 55 60
 Lys Glu Asp Thr Thr Asp Thr Gln Tyr Asp Gln Asp Phe Ser Lys Tyr
 65 70 75 80
 Thr Asn Asn Asn Phe Ile Asp Ser Ser Phe Tyr Asp Asp Lys Pro Asn
 10 85 90 95
 Met Phe Asp Phe Thr Leu Ser His Arg Thr Lys Ala Cys Leu Glu Ser
 100 105 110
 Val Lys Met Gly Val Lys Met Gly Thr Met Val Gly Gly Ile Phe Gly
 115 120 125
 15 Ser Leu Thr Gly Leu Tyr Ala Ser Phe Ala His Lys Asn Leu Phe Ile
 130 135 140
 Phe Pro Val Ser Val Ile Gly Gly Ala Val Ser Phe Gly Phe Phe Leu
 145 150 155 160
 Gly Cys Gly Met Ile Val Arg
 20 165

<210> 13

<211> 107

<212> PRT

25 <213> 稻瘟病菌 (*Magnaporthe grisea*)

<400> 13

Met Pro Pro Pro Pro Arg Gln Ala Gly His Gly Ala Ala Pro Ser Asn
 1 5 10 15
 30 Met Asp Lys Pro Val Gly Ala Ile Met Gly Phe Ile Tyr Gly Thr Val
 20 25 30
 Thr Val Phe Arg Gly Gly Ala Gly Pro Asn Gly Ile Met Arg Thr Ile
 35 40 45
 Gly Gln Tyr Met Leu Ala Ser Gly Thr Thr Phe Gly Phe Phe Met Gly
 35 50 55 60

Ile Gly Ser Val Ile Arg Ser Asp Ala Ser Pro Ile Ser Gln Gln Ala
65 70 75 80
Tyr Phe Gln Thr Arg Pro Arg Pro Leu Ile Met Ala Ser His Arg Ala
85 90 95
5 Phe Arg Pro Gln Gln Ser Thr Arg Arg Asn Asp
100 105

<210> 14
<211> 129
10 <212> PRT
<213> 脉孢菌 (*Neurospora crassa*)

<400> 14
Met Pro Pro Pro Pro Gln His Gly Gly Ala Val Gly Pro Ser Asn Phe
15 1 5 10 15
Asp Lys Phe Lys Met Gly Ala Met Met Gly Gly Ser Pro Ser Pro Lys
20 25 30
Asp Ser Ile Asp Leu Arg Ala Cys Pro Asp Val Arg His Ser Pro Gly
35 40 45
20 Thr Val Asn Ile Phe Arg Tyr Gly Ala Gly Pro Asn Gly Ile Met Arg
50 55 60
Thr Leu Gly Gln Tyr Met Leu Gly Ser Gly Ala Thr Phe Gly Arg Pro
65 70 75 80
Ser Ser Phe Phe Met Ser Ile Gly Ser Val Ile Arg Ser Asp Ser Ser
25 85 90 95
Pro Ile Val Ala Glu Ala Tyr Tyr Arg Ala Gln Arg Arg Pro Met Ile
100 105 110
Met Ala Ala Gln Ala Phe Arg Pro Ala Tyr Tyr Pro Thr Arg Arg Ser
115 120 125
30 Asp

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2004/001435

A. CLASSIFICATION OF SUBJECT MATTER

C07K14/47, C12N15/12, C12N15/63, C12N15/10, A61K38/17, A61P31/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K, C12N, A61K, A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, PAJ, EPODOC, CNPAT, CNKI, CA, GenBank

Keywords: Glycine-rich protein? (GRP?), glycine, peptide?, antibacteribal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|-----------------------|
| X | GenBank (www.ncbi.nlm.nih.gov), Accession No. AY028425, 19.MAR.2001 | 1-10 |
| P,X | GenBank (www.ncbi.nlm.nih.gov), Accession No. NP_542786, 27.Oct.2004 | 1-10 |
| A | University of veterinary medicine, Vol.42, No.1, 1998, Glinski Zd et al., "Novel antibacterial insect immune peptides and proteins", Pages 33-41, especially abstract | 1-14 |
| A | Journal of Experimental Hematology, Vol.10, No.3, 2002, ZHANG Yong et al., "Screening of Differentially Expressed Genes in the Mouse Hematopoietic Stromal Cells after Long Term Culture", Pages 177-182 | |

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

02.MAR.2005 (02.03.2005)

Date of mailing of the international search report

31-MAR-2005 (31-03-2005)

Name and mailing address of the ISA/CN

6 Xitucheng Rd., Jimen Bridge, Haidian District,
100088 Beijing, China

Authorized officer

ZHOU Xia

Telephone No. 86-10-62085078

Facsimile No. 86-10-62019451

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2004/001435

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See further information sheet PCT/ISA/210

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2004/001435

This International Searching Authority found multiple inventions in this international application, as follows:

Group 1: claims 1-15, relating to the glyrichins basing on the sequence of SEQ ID NO:1 from human and mouse;

Group 2: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:3 from *Daniorerio*;

Group 3: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:4 from *Anopheles gambiae*;

Group4: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:5 from *Drosophila melanogas*;

Group5: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:6 from *Caenorhabditis elegans*;

Group6: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:7 from *Caenorhabditis elegan*;

Group7: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:8 from *Schizosaccharomyces pombe*;

Group8: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:9 from *Sacchchromyces cerevisiae*;

Group9: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:10 from *Arabidopsis thaliana*;

Group10: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:11 from *Plasmodium falciparum 3D7*;

Group11: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:12 from

Group12: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:13 from *Magnaporthe grisea*;

Group13: claims 1, 3-6, 8-15, relating to the glyrichins basing on the sequence of SEQ ID NO:14 from *Neurospora crassa*;

Since the inventions of Group1-Group 13 related to the different glycine-rich proteins(GRPs) having the different amino acid sequences from special biology sources, these inventions did not share any of the technical features identified, the international application does not comply with the requirements of unity of invention as defined in PCT rule 13.

国际检索报告

国际申请号
PCT/CN2004/001435

A. 主题的分类

C07K14/47, C12N15/12, C12N15/63, C12N15/10, A61K38/17, A61P31/04

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

C07K, C12N, A61K, A61P

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

WPI, PAJ, EPODOC, CNPAT, CNKI, CA, GenBank

Glycine-rich protein? (GRP?), glycine, peptide?, antibacteribal

C. 相关文件

| 类 型* | 引用文件, 必要时, 指明相关段落 | 相关的权利要求 |
|------|---|---------|
| X | GenBank (www.ncbi.nlm.nih.gov) 登录号 AY028425, 2001 年 3 月 19 日 | 1-10 |
| P, X | GenBank (www.ncbi.nlm.nih.gov) 登录号 NP_542786, 2004 年 10 月 27 日 | 1-10 |
| A | University of veterinary medicine, 第 42 卷, 第 1 期, 1998 年, Glinski, Zd. 等, "Novel antibacterial insect immune peptides and proteins", 第 33-41 页, 特别是摘要 | 1-14 |
| A | 中国实验血液学杂志, 第 10 卷, 第 3 期, 2002 年出版, 张咏等, "长期培养后小鼠造血基质细胞差异表达基因的筛选", 第 177-182 页 | |

☐ 其余文件在 C 栏的续页中列出。

☐ 见同族专利附件。

* 引用文件的具体类型:

"A" 认为不特别相关的表示了现有技术一般状态的文件

"E" 在国际申请日的当天或之后公布的在先申请或专利

"L" 可能对优先权要求构成怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

"O" 涉及口头公开、使用、展览或其他方式公开的文件

"P" 公布日先于国际申请日但迟于所要求的优先权日的文件

"T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

"X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

"Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

"&" 同族专利的文件

国际检索实际完成的日期

02.3 月 2005 (02.03.2005)

国际检索报告邮寄日期

31. 3 月 2005 (31. 03. 2005)

中华人民共和国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路 6 号 100088

传真号: (86-10)62019451

授权官员

周霞



电话号码: (86-10)62085078

国际检索报告

国际申请号

PCT/CN2004/001435

第II栏 关于某些权利要求不能作为检索主题的意见(接第1页第2项)

按条约 17(2)(a)对某些权利要求未作国际检索报告的理由如下:

1. ☐ 权利要求:
因为它们涉及到不要求本国际检索单位进行检索的主题, 即:

2. ☐ 权利要求:
因为它们涉及到国际申请中不符合规定的要求的部分, 以致不能进行任何有意义的国际检索,
具体地说:

3. ☐ 权利要求:
因为它们是从属权利要求, 并且没有按照细则 6.4(a)第 2 句和第 3 句的要求撰写。

第II栏 关于缺乏发明单一性时的意见(接第1页第3项)

本国际检索单位在该国际申请中发现多项发明, 即:

参见 PCT/ISA/210 表(附加页)。

1. ☐ 由于申请人按时缴纳了被要求缴纳的全部附加检索费, 本国际检索报告针对全部可作检索的权利要求。
2. ☒ 由于无需付出有理由要求附加费的劳动即能对全部可检索的权利要求进行检索, 本国际检索单位未通知缴纳任何附加费。
3. ☐ 由于申请人仅按时缴纳了部分被要求缴纳的附加检索费, 本国际检索报告仅涉及已缴费的那些权利要求。具体地说, 是权利要求:
4. ☐ 申请人未按时缴纳被要求的附加检索费。因此, 本国际检索报告仅涉及权利要求中首次提及的发明; 包含该发明的权利要求是:

关于异议的说明: ☐ 申请人的异议书随附加检索费同时提交。

☐ 支付附加检索费时未提交异议书。

本申请中存在下述 13 组发明:

组 1: 权利要求 1-15, 涉及来源于人和小鼠的基于序列 1 的 glyrichin;

组 2: 权利要求 1、3-6、8-15, 涉及来源于斑马鱼 (*Danio rerio*) 的基于序列 3 的 glyrichin;

组 3: 权利要求 1、3-6、8-15, 涉及来源于按蚊 (*Anopheles gambiae*) 的基于序列 4 的 glyrichin;

组 4: 权利要求 1、3-6、8-15, 涉及来源于果蝇 (*Drosophila melanogaster*) 的基于序列 5 的 glyrichin;

组 5: 权利要求 1、3-6、8-15, 涉及来源于线虫 (*Caenorhabditis elegans*) 的基于序列 6 的 glyrichin;

组 6: 权利要求 1、3-6、8-15, 涉及来源于线虫 (*Caenorhabditis elegans*) 的基于序列 7 的 glyrichin;

组 7: 权利要求 1、3-6、8-15, 涉及来源于芽殖酵母 (*Schizosaccharomyces pombe*) 的基于序列 8 的 glyrichin;

组 8: 权利要求 1、3-6、8-15, 涉及来源于酿酒酵母 (*Saccharomyces cerevisiae*) 的基于序列 9 的 glyrichin;

组 9: 权利要求 1、3-6、8-15, 涉及来源于拟南芥 (*Arabidopsis thaliana*) 的基于序列 10 的 glyrichin;

组 10: 权利要求 1、3-6、8-15, 涉及来源于疟原虫 (*Plasmodium falciparum* 3D7) 的基于序列 11 的 glyrichin;

组 11: 权利要求 1、3-6、8-15, 涉及来源于疟原虫 (*Plasmodium yoelii yoelii*) 的基于序列 12 的 glyrichin;

组 12: 权利要求 1、3-6、8-15, 涉及来源于稻瘟病菌 (*Magnaporthe grisea*) 的基于序列 13 的 glyrichin;

组 13: 权利要求 1、3-6、8-15, 涉及来源于脉孢菌 (*Neurospora crassa*) 的基于序列 14 的 glyrichin;

由于上述 13 组富含甘氨酸蛋白 (Glyrichin) 的组成序列不同, 故它们之间不存在相同或相应的特定技术特征, 故上述 13 组发明之间不具备单一性, 不符合 PCT 实施细则第 13 条的规定。